The Volatility of the Instantaneous Spot Interest Rate Implied by Arbitrage Pricing - A Dynamic Bayesian Approach

Ramaprasad Bhar* Carl Chiarella† Hing Hung‡
Wolfgang J. Runggaldier§

January 30, 2004

$[-Carl * runggaldier - BCR1 * new7 - BCR1 * new5 - BCR1 * new5 * BAYESversion1.tex]$

Abstract

This paper considers the estimation of the volatility of the instantaneous short interest rate from a new perspective. Rather than using discretely compounded market rates as a proxy for the instantaneous short rate of interest, we derive a relationship between observed LIBOR rates and certain unobserved instantaneous forward rates. We determine the stochastic dynamics for these rates under the risk-neutral measure and propose a filtering estimation algorithm for a time-discretised version of the resulting interest rate dynamics based on dynamic Bayesian updating. The method is applied to US Treasury rates of various maturities and is found to give a reasonable model fit.

*School of Banking and Finance, The University of New South Wales, Sydney, r.bhar@unsw.edu.au

†School of Finance and Economics, University of Technology, Sydney, Australia, E-mail: Carl.Chiarella@uts.edu.au

‡School of Finance and Economics, University of Technology, Sydney, Australia, E-mail: hing.hung@uts.edu.au

§Dipartimento di Matematica Pura ed Applicata, Universitá di Padova, E-mail:runggal@math.unipd.it

1 Introduction

Literature on estimation in models of the instantaneous spot rate of interest has burgeoned since the seminal contribution of Chan et al. (1992) (henceforth CKLS). The sustained interest in this topic is due to the great deal of activity, both amongst academics and practitioners, in pricing interest rate related securities.

CKLS applied the generalised method of moments (GMM) to estimate the parameters of a one factor model of the instantaneous spot rate of interest that is mean-reverting in the drift term and has a diffusion term that is of constant elasticity in the spot rate. Their estimate for US data of about 1.5 for the elasticity in the diffusion term provoked much discussion as it was much higher than the value of 0.5 for the popular Cox, Ingersoll and Ross (1985) model.

Subsequent contributions either extended the basic CKLS formulation and/or considered alternative estimation procedures.

Longstaff and Schwartz (1992), Brenner et al. (1996), Andersen and Lund (1997) and Koedijk et al. (1997) in various ways add volatility dynamics to the model for interest rate dynamics. Sun (2003) considers a quite general specification allowing for a non-linear drift as well as ARCH-type stochastic volatility.

As far as estimation methodology is concerned, GMM has remained the work-horse for most of the empirical studies cited. However Nowman (1997, 1998) applied the Gaussian estimation techniques developed by Bergstrom (1990) for continuous time stochastic differential equations. In contrast to GMM the Gaussian estimation methodology has the advantage of producing an exact maximum likelihood estimator. Episcopos (2000) subsequently applied this methodology to estimate the parameters of the CKLS specification for the short-term interest rate for a number of countries. Interestingly he obtained estimates for the elasticity in the diffusion term that are much lower than those obtained by CKLS. The fact that two well established estimation methodologies can yield widely differing parameter estimates suggests a need to look at the estimation issue from a different perspective and that is one of the motivations for the current paper.

Irrespective of the estimation methodology one employs, another significant issue relates to what data is used to proxy the instantaneous spot rate of interest that itself is not an observed quantity. Proxy variables that have been used include US one-month Treasury bill rates (CKLS) and one-month interbank rates (Episcopos). It seems strange that the literature has not developed in the direction of deriving and including in the estimation procedure, the stochastic differential equations that the assumed dynamics of the instantaneous spot rate imply for these market observed rates. Certainly Chapman et al. (1999) have provided some evidence that use of the indicated proxy variables may not induce a great deal of error. However given that the choice of appropriate estimation procedures is not yet an entirely settled issue, it would seem useful to establish a framework that removes entirely any potential for errors or biases due to choice of the proxy variable. The current paper provides such a framework.

In this paper we use the framework of Heath-Jarrow-Morton (1992) (henceforth HJM) to model the dynamics of the interest rate market. The starting point of HJM is a specification of the dynamics of the forward rate to any general maturity. We specify a forward rate volatility function that yields the same volatility function for the instantaneous spot rate of interest considered in the earlier cited literature. An important difference is that the dynamics of the interest rate processes occur under the risk-neutral measure. Under this measure the HJM procedures enable us to obtain the dynamics of pure discount bond prices. These can in turn be related to the discretely compounded LIBOR rates. This link then enables us to determine the dynamics for LIBOR rates. It turns out that the dynamics of the LIBOR rate, the instantaneous spot rate of interest and another instantaneous forward rate evolve simultaneously under the risk-neutral measure. The link between pure discount bond prices and LIBOR rates mean that these rates can be regarded as observable under the risk-neutral measure, whilst the other two instantaneous rates just referred to are not observable. We are thus dealing with a partially observed stochastic dynamic system whose estimation may be undertaken by the use of non-linear filtering methods. Here we develop a dynamic Bayesian updating algorithm similar to the one proposed in Chiarella, Pasquali and Ruggaldier (2001). The basic approach proposed here has been applied to a much simpler (and approximate) representation of discrete tenor interest rate dynamics in Bhar, Chiarella and Runggaldier (2002).

The plan of the paper is as follows: in section 2 we derive the stochastic dynamic system followed by the instantaneous spot rate and discretely compounded LIBOR rates. Since the data are observed in discrete time, in section 3 we outline the way in which the continuous time stochastic differential equation system is discretised. In section 4 we outline the way in which the dynamic Bayesian updating algorithm is applied to the estimation problem. In section 5 we discuss implementation issues and apply the algorithm to some U.S. data. Section 6 concludes and makes suggestions for future research. Detailed technical derivations are relegated to the appendix.

2 The Dynamics of LIBOR Rates Implied by HJM Bond Prices

We use the Brace and Musiela (1994) (henceforth BM) parameterisation of the HJM model, which is in terms of $r(t,x)(x \geq 0)$the x-period instantaneous forward rate at time t for maturity $(t+x)$. Under the risk-neutral measure $\widetilde{P}$ this rate satisfies the stochastic integral equation

$$r(t,x) = r(0,t+x) + \int_0^t \sigma(s,t+x)\overline{\sigma}(s,t+x)ds + \int_0^t \sigma(s,t+x)d\widetilde{w}(s), \tag{1}$$

where $r(0,x)$ is the initial forward curve, $\widetilde{w}$ is a Wiener process under $\widetilde{\mathbb{P}}$ and $\sigma(t,x)$ is the instantaneous forward rate volatility function that could (and in our application will)depend on certain instantaneous forward rates. In equation (1)

$$\overline{\sigma}(s,t+x) = \int_s^{t+x} \sigma(s,u)du. \tag{2}$$

It is important to stress that even though we use the BM parameterisation for the forward rate dynamics, we use the notation for the volatility function as in HJM in that $\sigma(t,x)$ refers to the forward rate volatility at time t applicable for time $x(\geqslant t)$. This is in contrast to BM who use the volatility function $\tau(t,x)$ to denote the forward rate volatility at time t applicable for time $t+x$. Of course these two different specifications of the forward rate volatility function are related via

$$\tau(t,x) = \sigma(t, t+x)$$

and one may work with either specification. For our application it turns out to be more convenient to use $\sigma(t,x)$.

In this notation the instantaneous spot rate of interest $r(t)$ is given by

$$r(t) = r(t,0) \tag{3}$$

and satisfies the stochastic integral equation

$$r(t) = r(0,t) + \int_o^t \sigma(v,t)\overline{\sigma}(v,t)dv + \int_0^t \sigma(v,t)d\widetilde{w}(v). \tag{4}$$

The price at time t of a $(t+x)$- maturity zero coupon bond is related to $r(t,x)$ by

$$b(t,x) = exp\left(-\int_0^x r(t,u)du\right). \tag{5}$$

Next we relate the x-period LIBOR rate to the bond price $b(t,x)$. We then derive the relationship between the bond price and the underlying state variables (a set of discrete tenor forward rates) upon which the forward rate volatility function depends. The dynamics of these state variables determine the evolution of the forward curve.

Consider a time period $(0,t)$ over which we have a set of observations of the x-period LIBOR rate, that we denote $L_x(t)$.This is an annualised rate at which $\$1$ invested at time t compounds simply to become $\$(1+xL_x(t))$ at time ($t+x$).

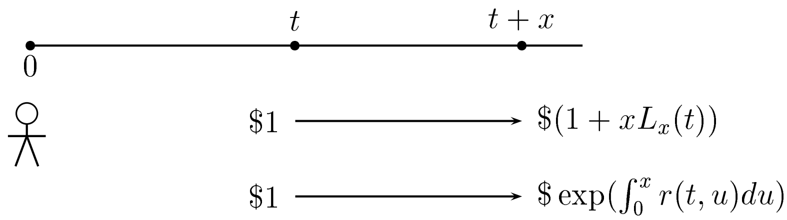


Figure 1: The LIBOR Rate $L_x(t)$

The LIBOR rate $L_x(t)$ is related to the continuously compounded Brace-Musiela forward rate by (see figure (1))

$$1 + xL_x(t) = exp(\int_o^x r(t,u)du). \tag{6}$$

From equations (6) and (5) we deduce the relationship between the LIBOR rate and the bond price, viz

$$L_x(t) = \frac{1}{x}(\frac{1}{b(t,x)} - 1). \tag{7}$$

However it turns out to be more convenient to work in terms of the quantity

$$l_x(t) = L_x(t) + \frac{1}{x}, \tag{8}$$

which is related to $b(t,x)$ via

$$l_x(t) = \frac{1}{xb(t,x)}. \tag{9}$$

We consider the volatility function of the general form

$$\sigma(t,u) = g(r(t,x_1),\cdots,r(t,x_n))e^{-\lambda(u-t)}, \tag{10}$$

where $\vec{r}(t,\cdot) \equiv [r(t,x_1),\cdots,r(t,x_n)]$ is a vector of discrete tenor forward rates chosen in the belief that these particular maturities most affect the evolution of the forward curve e.g. perhaps they correspond to the most liquid maturities. In our subsequent application we shall specialise (10) to the case where $g(\cdot)$ depends on just one argument and has the particular form

$$g(r) = \sigma_0(\min\{\varepsilon^{-1}, \max[|r|,\varepsilon]\})^{\delta} \tag{11}$$

where $\delta > 0$ and σ_0 are parameters to be estimated and $\varepsilon > 0$ is a given, arbitrarily small constant.

This representation is consistent with the earlier cited empirical literature that concentrates merely on dependence on the instantaneous short rate. We use $\min\{\varepsilon^{-1}, \max[|r|,\varepsilon]\}$ to the prevent the volatility from becoming either zero or infinite.

Thus

$$\sigma(t,u) = g(r(t))e^{-\lambda(u-t)}. \tag{12}$$

Subsequent applications could allow for dependence on a number of different tenor forward rates.

Chiarella and Kwon (2003) show that with the specification (12) the bond price may be expressed as a deterministic combination of two discrete tenor forward rates $r(t,x_1)$, $r(t,x_2)$ whose tenors may be chosen arbitrarily. The relevant details are summarized in Appendix 2 from equation (72) of which we have

$$b(t,x) = exp(-[\bar{b}_o(t,x) + \sum_{i=1}^{2} \bar{b}_i(t,x)r(t,x_i)]), \tag{13}$$

where the $\bar{b}_i(t,x)(i=0,1,2)$ are defined in general by equations (70) and (73) and evaluated for the specific volatility function (12) in equations (84)-(86). The stochastic differential equations followed by the $r(t,x_k)(k=1,2,\cdots,n)$ are given by equations (75) of appendix 2, namely [1]

$$
\begin{aligned}
dr(t,x_k) = {} & [b_0^{'}(t,x_k) + \sum_{i=1}^{2} b_i^{'}(t,x_k) r(t,x_i) + \sigma(t,t+x_k)\bar{\sigma}(t,t+x_k)]dt \\
& + e^{-\lambda x_k} g(r(t,x_1),...,r(t,x_n)) d\widetilde{w}(t), \qquad (k=1,2,...,n).
\end{aligned} \tag{14}
$$

Keeping in mind that our aim is to estimate the parameters $(\sigma_0,\delta,\lambda)$ used to specify the particular volatility structure (12), we use the foregoing term structure dynamics as follows. First we treat equation (9) for $l_x(t)$ (with $b(t,x)$ given by (13)) as the observation equation, with underlying unobserved state variables $r(t,x_1), r(t,x_2),\cdots, r(t,x_n)$ being driven by the system (14). Note that here we have set things up in such a way that the $r(t,x_1), r(t,x_2)$ appearing in (13) are the first two elements of the vector $\vec{r}(t,\cdot)$ upon which the volatility function is dependent. It should be stressed that this choice is somewhat arbitrary and any two elements of $\vec{r}(t,\cdot)$ might have been chosen. Indeed it is possible to use two discrete tenor forward rates not belonging to $\vec{r}(t,\cdot)$, in which case an additional two stochastic differential equations for their dynamics would have to be appended to the system (14). The particular choices made in this regard are implementation issues.

Turning to our particular implementation with the volatility function (12), this fits into the general structure of equation (10) by setting $n=1$ and $x_1=0$ so that

$$\vec{r}(t,\cdot) = r(t,0) = r(t).$$

In equation (13) we also set $r(t,x_1)=r(t,0)=r(t)$ and leave $r(t,x_2)$ as some arbitrary tenor discrete forward rate. The dynamics for $r(t,x_2)$ will append an additional stochastic differential equation to the one to which the system (14)reduces in this case.

To summarise, the expression for $b(t,x)$ will be given by

$$b(t,x) = exp(-\left[\bar{b}_0(t,x) + \bar{b}_1(t,x)r(t) + \bar{b}_2(t,x)r(t,x_2)\right]), \tag{15}$$

where the dynamics for $r(t)$ and $r(t,x_2)$ are given by

$$
\begin{aligned}
dr(t) = & [b_0'(t,0) + b_1'(t,0)r(t) + b_2'(t,0)r(t,x_2)]dt \\
& + g(r(t))d\widetilde{w}(t),
\end{aligned}
$$

and

$$
\begin{aligned}
dr(t,x_2) = & [b_0'(t,x_2) + b_1'(t,x_2)r(t) + b_2'(t,x_2)r(t,x_2) \\
& + \sigma(t,t+x_2)\bar{\sigma}(t,t+x_2)]\,dt \\
& + e^{-\lambda x_2} g(r(t))d\widetilde{w}(t).
\end{aligned}
$$

As we have stated, the choice of x_2 is arbitrary, for an initial implementation we choose x_2 to be the same as the tenor x of the observed LIBOR rates.

[1]Note that $b_i'(t,x) \equiv \frac{\partial}{\partial x} b_i(t,x)$ and the precise expressions are given in equations (81)-(83).

Here we should stress that the driving dynamics (14) are under the risk neutral measure $\widetilde{\mathbb{P}}$. However the LIBOR rates are observed under the real world measure $\mathbb{P}$. To convert the dynamic (14) to the dynamics under $\mathbb{P}$ we would have to introduce the market price of interest rate risk. However if we assume (as we shall) that the market price of risk is at most a time deterministic function (and so in particular not stochastic) then the diffusion of the underlying process will be the same under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. Of course the drifts will differ but we are not concerned in this paper with estimating the drift term. If the market price of risk were stochastic, either through dependence on some of the state variables, or because it follows some independent stochastic differential equation then we would need to consider the dynamics and estimation procedure under the historical measure. The Bayesian updated algorithm (appropriately modified) to be described below could still be applied to the resulting stochastic dynamical system.

3 State Space Form of the Model

Summarizing the results of the previous section, we shall take as our partially observable system the (unobservable) instantaneous x-period forward rate $r(t,x)$, and instantaneous spot rate $r(t,0) \equiv r(t)$. The set of stochastic differential equations for the state may be written

$$\begin{aligned} dr(t,x) &= R_x(r(t,x), r(t,0))dt + e^{-\lambda x} g(r(t,0))d\widetilde{w}, && (16)\\ dr(t,0) &= R(r(t,x), r(t,0))dt + g(r(t,0))d\widetilde{w}, && (17) \end{aligned}$$

where we set

$$\begin{aligned} R_x(r(t,x), r(t,0)) =& b_0'(t,x) + b_1'(t,x)r(t,0) + b_2'(t,x)r(t,x) \\ &+ \sigma(t,t+x)\bar{\sigma}(t,t+x) \end{aligned} \tag{18}$$

and

$$R(r(t,x), r(t,0)) = b_0'(t,0) + b_1'(t,0)r(t,0) + b_2'(t,0)r(t,x). \tag{19}$$

Equations (16), (17) are the state transition equations for the unobserved state variables, $r(t,x)$, $r(t,0)$.

We shall use $S := (r(t,0), r(t,x))^T$ to denote the state vector. This enables us to write the state transition equations (16) as

$$dS(t) = F(S(t);\theta)dt + V(S(t);\theta)d\widetilde{w}, \tag{20}$$

where

$$\begin{aligned} F(S;\theta) &= (R(r(t,x), r(t,0)), R_x(r(t,x), r(t,0)))^T, && (21)\\ V(S;\theta) &= (g(r(t,0)), e^{-\lambda x} g(r(t,0)))^T, && (22) \end{aligned}$$

and θ is the (to be estimated) parameter vector

$$\theta = (\sigma_0,\ \delta,\ \lambda).$$

Financial implementations of estimation methodologies are usually carried out in a discrete time setting as data are observed discretely. Thus we discretise (16), (17) using the Euler-Maruyama scheme to become

$$
\begin{aligned}
r(k+1,x) - r(k,x) &= R_x(r(k,x), r(k,0))\Delta t + e^{-\lambda x} g(r(k,0))\Delta W_k \\
r(k+1,0) - r(k,0) &= R(r(k,x), r(k,0))\Delta t + g(r(k,0))\Delta W_k
\end{aligned}
\tag{23}
$$

where, Δt denotes the time steps and see (18), (19),

$$
\begin{aligned}
R_x(r(k,x), r(k,0)) &= b_0'(k\Delta t, x) + b_1'(k\Delta t, x) r(k,0) + b_2'(k\Delta t, x) r(k,x) + \sigma(k, k+x)\bar{\sigma}(k, k+x) \\
R(r(k,x), r(k,0)) &= b_0'(k\Delta t, 0) + b_1'(k\Delta t, 0) r(k,0) + b_2'(k\Delta t, 0) r(k,x).
\end{aligned}
\tag{24}
$$

Equation (23) can be synthesized as

$$S_{k+1} - S_k = F_k(S_k;\theta)\Delta t + V_k(S_k;\theta)\Delta W_k \tag{25}$$

with $F_k(\cdot)$ and $V_k(\cdot)$ corresponding to (21) and (22) respectively and where $\Delta W_k \sim \mathcal{N}(0, \Delta t)$. We have used the shorthand notation $r(k,x)$ to represent $r(k\Delta, x)$ and, similarly, with other quantities such as S_k standing for $S_{k\Delta t}$.

We take as the observations in our system the LIBOR rates $l_x(t)$. From equations (9) and (15), and using the approximation $\ln l_x(t) \simeq l_x(t) - 1$ we have that $l_x(t)$ is related to the state variables by

$$l_x(t) = 1 - \ln x + \bar{b}_0(t,x) + \bar{b}_1(t,x) r(t) + \bar{b}_2(t,x) r(t,x) \tag{26}$$

Using Y_k to denote $l_x(k\Delta t)$, the observation equation becomes

$$Y_k = C_k S_k + \bar{b}_0(k\Delta t, x) + 1 - \ln x + q_k \eta_k, \tag{27}$$

where

$$C_k = (\bar{b}_1(k\Delta t, x), \bar{b}_2(k\Delta t, x)). \tag{28}$$

We have assumed in (27) the existence of an observation noise term $q_k\eta_k$, where $\eta_k \sim \mathcal{N}(0,1)$ is serially uncorrelated and independent of the ε_k. The strength of the observation noise, q_k, would reflect features (such as bid-ask spread) of the LIBOR market. In order to express the observation equation (27) in standard form we define the noise term

$$\lambda_k = q_k \eta_k, \tag{29}$$

so that

$$\lambda_k \sim N(0, \Lambda_k), \tag{30}$$

where

$$\Lambda_k = q_k^2. \tag{31}$$

With this notation the observation equation (27) may be written

$$Y_k = C_k S_k + \bar{b}_0(k\Delta t, x) + 1 - \ln x + \lambda_k. \tag{32}$$

4 The Dynamic Bayesian Updating Algorithm

From (25) we obtain for the conditional distribution of S_{k+1}, given S_k, the Gaussian distribution

$$p(S_{k+1}|S_k) \sim \mathcal{N}(S_{k+1}; S_k + F_k(S_k, \theta t)\Delta t, V_k(S_k, \theta t)V_k'(S_k, 0)\Delta t) \tag{33}$$

where we use the notation $\mathcal{N}(X; m, \Sigma)$ to denote a Gaussian random variable X with mean m and covariance matrix Σ. More specifically, we have

$$V_k(S_k, \theta)V_k'(S_k, \theta)\Delta = \begin{pmatrix} g^2(r(k,0)\Delta t) & e^{-\lambda x} g^2(r(k,0)\Delta t) \\ e^{-\lambda x} g^2(r(k,0)\Delta t) & e^{-2\lambda x} g^2(r(k,0))\Delta t \end{pmatrix} \tag{34}$$

which is immediately seen to be singular (the conditional correlation among the two components is equal to 1 and their joint distribution degenerates). The two components of the state vector are in fact linearly dependent and one can write

$$\begin{aligned} r(k+1, x) &= e^{-\lambda x} r(k+1, 0) + [(r(k,x) + R_x(r(k,x), r(k,0)) - e^{-\lambda x}(r(k,0) + R(r(k,x), r(k,0)))] \\ &= \alpha r(k+1, 0) + \beta(r(k,x), r(k,0)) \end{aligned} \tag{35}$$

thereby implicitly defining the constant α and the function $\beta(\cdot)$ that (see (24) with (11), (12) and (2)) is uniformly continuous in its arguments.

Putting $S_k = (S_k^1, S_k^2)$ where $S_k^1 = r(k, 0)$, $S_k^2 = r(k, x)$, and making use of (35) we may also rewrite (25) as

$$\begin{aligned} S_{k+1}^1 &= S_k^1 + F_k^1(S_k, \theta)\Delta t + V_k^1(S_k, \theta)\Delta W_k, \\ S_{k+1}^2 &= \alpha S_{k+1}^1 + \beta(S_k), \end{aligned} \tag{36}$$

where F_k^1 and V_k^1 are the first components of the 2-vectors F_k and V_k in (25) respectively. Below we shall also consider the conditional distribution of S_{k+1}^1, given S_k, that is induced by (36) namely

$$p_\theta(S_{k+1}^1|S_k) \sim \mathcal{N}(S_{k+1}^1; S_k^1 + F_k^1(S_k, \theta)\Delta t, (V_k^1(S_k, \theta))^2\Delta t) \tag{37}$$

and that admits a density with respect to the Lebesgue measure on $\mathbb{R}^1$ (notice that $p_\theta(S_{k+1}|S_k)$, being degenerate, does not admit a density with respect to the Lebesgue measure on $\mathbb{R}^2$).

On the other hand, from (32) we obtain the conditional distribution of Y_k given S_k and θ as

$$p(Y_k|S_k, \theta) = \frac{1}{\sqrt{2\pi\Lambda_k}} \exp\left\{-\frac{1}{2\Lambda_k}\left[Y_k - C_k S_k - \bar{b}_0(k\Delta t, x) - 1 + \ln x\right]^2\right\} \tag{38}$$

Using the representation (36) of the dynamics (25) we may also consider the distribution of Y_k given S_k^1, S_{k-1} and θ namely

$$\begin{aligned} \bar{p}\left(Y_k|S_k^1, S_{k-1}; \theta\right) = \frac{1}{\sqrt{2\pi\Lambda_k}} \exp\left\{-\frac{1}{2\Lambda_k}\left[Y_k - \bar{b}_1(k\Delta t, x)S_k^1 - \bar{b}_2(k\Delta t, x)\alpha S_k^1\right.\right. \\ \left.\left. - \bar{b}_2(k\Delta t, x)\beta(S_{k-1}) - \bar{b}_0(k\Delta t, x) - 1 + \ln x\right]^2\right\}. \end{aligned} \tag{39}$$

We are interested in the conditional joint distribution $p(S_k, \theta|y^k)$ of S_k and θ in the generic period $k = k\Delta t$, given the observations $y^k = (y_1, ..., y_k)$. Notice that, since the two components of S_k are linearly dependent, $p(S_k, \theta|y^n)$ does in general not have a density with respect to the Lebesgue measure on $\mathbb{R}^2$.

However, it still satisfies the recursive Bayes formula, namely

$$p(S_{k+1}, \theta|y^{k+1}) \propto p(Y_{k+1}|S_{k+1}, \theta) \int p_\theta(S_{k+1}|S_k) dp(S_k, \theta|y^k) \tag{40}$$

with initial condition that we choose of the form $p(S_0, \theta) = p(S_0)p(\theta)$ (independence of S_0 and θ) and where $\propto$ denotes "proportional to".

Recall that θ is supposed from the outset to take only a finite number of values and so in the Bayesian context it is considered as a discrete random variable.

Since S_k takes a continuum of possible values (its dynamics are driven by the Gaussian ΔW_k), to actually compute the recursion (40) we discretize the values of S_k (for the convengence of the ensuing approximation see Proposition 4.1 below).

For this purpose, given a step-size $\delta > 0$ and an integer H, consider the square in $\mathbb{R}^2$ given by $(-H\delta, H\delta] \times (-H\delta, H\delta]$ and its partition (grid) into $4H^2$ squares, the generic l-th $(1 \leq l \leq 4H^2)$ of which is derived as follows: let

$$l = (2H)h + k + 1$$

where $0 \leq h < 2H$, $0 \leq k < 2H$, then

$$\begin{aligned} R^l &= ((-H+k)\delta, (-H+k+1)\delta) \times ((-H+h)\delta, (-H+h+1)\delta)) \\ &:= (\alpha_1^l, \beta_1^l] \times (\alpha_2^l, \beta_2^l] \end{aligned} \tag{41}$$

thereby implicitly defining $\alpha_1^l, \beta_1^l, \alpha_2^l, \beta_2^l$. In each of the $4H^2$ square R^l pick a representative element, e.g. its middle point, namely

$$\left(\left(-H + k + \frac{1}{2}\right)\delta, \left(-H + h + \frac{1}{2}\right)\delta, \right) \tag{42}$$

if $l = (2H)h + k + 1$ (alternative choices for representative elements are equally valid).

In addition to the squares R^l for $l \leq 4H^2$ that form a partition of $(-H\delta, H\delta] \times (-H\delta, H\delta]$, consider further 8 subsets of $\mathbb{R}^2$, denoted also by R^l, and that for $l = 4H^2 + 1, ..., 4H^2 + 8$ are given by (on the right of each set there appears a possible choice for its representative element)

$$
\begin{aligned}
R^{4H^2+1} &= (-\infty, -H\delta] \times (-\infty, -H\delta], &\qquad& \left((-H-\frac{1}{2})\delta, (-H-\frac{1}{2})\delta\right) \\
R^{4H^2+2} &= (-H\delta, H\delta] \times (-\infty, -H\delta], && \left(0, (-H-\frac{1}{2})\delta\right) \\
R^{4H^2+3} &= (H\delta, +\infty] \times (-\infty, -H\delta], && \left((H+\frac{1}{2})\delta, (-H-\frac{1}{2}\delta)\right) \\
R^{4H^2+4} &= (-\infty, -H\delta] \times (-H\delta, H\delta], && \left((-H-\frac{1}{2})\delta, 0\right) \\
R^{4H^2+5} &= (H\delta, +\infty] \times (-H\delta, H\delta], && \left((H+\frac{1}{2})\delta, 0\right) \\
R^{4H^2+6} &= (-\infty, -H\delta] \times (H\delta, +\infty], && \left((-H-\frac{1}{2})\delta, (H+\frac{1}{2})\delta\right) \\
R^{4H^2+7} &= (-H\delta, H\delta] \times (H\delta, +\infty], && \left(0, (H+\frac{1}{2})\delta\right) \\
R^{4H^2+8} &= (H\delta, \infty] \times (H\delta, +\infty]. && \left((H+\frac{1}{2})\delta, (H+\frac{1}{2})\delta\right)
\end{aligned}
$$

We shall use the notation $R^l = (\alpha_1^l, \beta_1^l] \times (\alpha_2^l, \beta_2^l)$ also for $l = 4H^2+1, ..., 4H^2+8$. Notice that $R^l(l = 1, ...4H^2+8)$ forms now a partition of all of $\mathbb{R}^2$ and we shall denote the chosen representative element of R^l by $\underline{r}^l = (r_1^l, r_2^l)$.

Given the discrete-time, continuous state Markov chain S_k with transition kernel $p_\theta(S_{k+1}|S_k)$, consider next the discrete-time finite state Markov chain $\bar{S}_k$, induced by S_k and having state space $\underline{r}^l(l = 1, ..., 4H^2+8)$. Denote by $P_{[i,h]}^{(k,\theta)}$ the generic (i, h)-th element $(i, h = 1, ...4H^2+8)$ of the corresponding transition probability matrix in period k. We then have, using the explicit representation (23) of the state transition equation (25).

$$
\begin{aligned}
P_{[i,h]}^{(k,\theta)} &= p_\theta\left\{\bar{S}_{k+1}=\underline{r}^h|\bar{S}_k=\underline{r}^i\right\} = P_\theta\left\{\left(S_{k+1}^1,S_{k+1}^2\right)\in R^h|(r_1^i,r_2^i)\right\}\\
&= P_\theta\{(r(k+1,x),r(k+1,0))\in R^h|(r(k,x),r(k,0))=(r_1^i,r_2^i)\}\\
&= P_\theta\left\{\alpha_1^h\le r(k+1,x)<\beta_1^h,\alpha_2^h\le r(k+1,0)<\beta_2^h|r(k,x)=r_1^i,r(k,0)=r_2^i\right\}\\
&= P_\theta\left\{\frac{\alpha_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)}\le\Delta W_k\le\frac{\beta_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)}\right.,\\
&\qquad\left.\frac{\alpha_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\le\Delta W_k\le\frac{\beta_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\right\}\\
&= P_\theta\left\{\max\left[\frac{\alpha_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)},\frac{\alpha_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\right]\le\Delta W_k\right.\\
&\qquad\left.\le\min\left[\frac{\beta_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)},\frac{\beta_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\right]\right\}\\
&= \Phi\left(\frac{1}{\sqrt{\Delta}}\min\left[\frac{\beta_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)},\frac{\beta_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\right]\right)\\
&\qquad-\Phi\left(\frac{1}{\sqrt{\Delta}}\max\left[\frac{\alpha_1^h-r_1^i-R_x(r_1^i,r_2^i)\Delta}{e^{-\lambda x}g(r_2^i)},\frac{\alpha_2^h-r_2^i-R(r_1^i,r_2^i)\Delta}{g(r_2^i)}\right]\right)
\end{aligned}
\tag{43}
$$

where $\Phi(\cdot)$ is the cummulative standard Gaussian distribution function.

The recursive Bayes' formula that corresponds to (40) for the discretized chain $\bar{S}_k$ becomes then

$$
\bar{p}\left(\bar{S}_{k+1}=\underline{r}^h,\theta|y^{k+1}\right)\propto p\left(Y_{k+1}|\bar{S}_{k+1}=\underline{r}^h,\theta\right)\sum_{i=1}^{4H^2+8}P_{[i,h]}^{(k,\theta)}\bar{p}(\bar{S}_k=\underline{r}^i,\theta|y^n) \tag{44}
$$

with initial condition $\bar{p}(S_0=\underline{r}^i,\theta)=P\{S_0\in R^i\}p_0(\theta)$ and the normalizing factor is the inverse of

$$
\sum_{h=1}^{4H^2+8}p(Y_{k+1}|\bar{S}_{k+1}=\underline{r}^h,\theta)\cdot\sum_{i=1}^{4H^2+8}P_{[i,h]}^{(k,\theta)}\cdot\bar{p}(\bar{S}_k=\underline{r}^i,\theta|y^k). \tag{45}
$$

Formula (44) can be written in matrix form as

$$
\bar{p}_{k+1}^\theta\propto L_{k+1}^\theta(P^{(k,\theta)})^T\bar{p}_k^\theta \tag{46}
$$

where $\bar{p}_k^\theta$ is the $(4H^2+8)$-vector with entries

$$
\bar{p}(\bar{S}_k=\underline{r}^i;\theta|y^k),\quad (i=1,\cdots,4H^2+8)
$$

$P^{(k,\theta)}$ is the matrix with elements $P_{[i,h]}^{(k,\theta)}$ $(i,h=1,\cdots,4H^2+8)$ and L_{k+1}^θ is the $(4H^2+8)\times(4H^2+8)$-diagonal matrix with entry in position h $(h=1,\cdots,4H^2+8)$ given by (see (38))

$$
\begin{aligned}
&p(Y_{k+1}|\bar{S}_{k+1}=\underline{r}^h,\theta)=\frac{1}{\sqrt{2\Lambda_{k+1}}}\cdot\\
&\exp\{-\frac{1}{2\Lambda_{k+1}}(Y_{k+1}-\bar{b}_2((k+1)\Delta t,x)r_1^h-\bar{b}_1((k+1)\Delta t,x)r_2^h-\bar{b}_0((k+1)\Delta t,x)+1-\ln x)^2\}.
\end{aligned}
\tag{47}
$$

From the joint conditional distribution $\bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k)$, $(h = 1, \cdots, 4H^2 + 8)$ one can obtain the marginal conditional distributions

$$\bar{p}(\bar{S}_k = \underline{r}^h | y^k) = \sum_{\theta} \bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k) \tag{48}$$

$$\bar{p}(\theta | y^k) = \sum_{h=1}^{4H^2+8} \bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k) \tag{49}$$

of $\bar{S}_k$ and θ respectively.

Combining (48) with (44) and (45) we obtain the explicit expression

$$\bar{p}(\bar{S}_k = \underline{r}^h | y^k) = \frac{\sum_{\theta} p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) \sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)} \cdot \bar{p}(\bar{S}_k = i, \theta | y^k)}{\sum_{\theta} \sum_{h=1}^{4H^2+8} p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) \sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)} \bar{p}(\bar{S}_k = i, \theta | y^k)} \tag{50}$$

and, analogously, for $\bar{p}(\theta | y^k)$.

We next show that the discretization introduced above to make the recursion (40) computable is meaningful by showing that the approximate conditional distributions computed via (44) converge in a suitable weak sense to the original conditional distribution corresponding to (40). Since θ is discrete already from the outset, it suffices that we consider the convergence of the conditional distributions for each fixed value of θ. We have in fact the following

Proposition 4.1 *(Weak convergence of conditional distributions)*

Given any uniformly continuous and bounded function $F(S)$ on $\mathbb{R}^2$, we have for any period k, for any sequence of observations y^k and for any of the finite values of θ

$$\lim_{\substack{H \to \infty \\ \delta \to 0}} \sum_{h=1}^{4H^2+8} F(\underline{r}^h) \bar{p}(\underline{r}^h, \theta | y^k) = \int F(S_k) dp(S_k, \theta | y^k) dS_k \tag{51}$$

where $\bar{p}(\underline{r}^h, \theta | y^k)$ is as in (44) and $p(S_k, \theta | y^k)$ as in (40).

<u>Proof:</u> See Appendix 2.

From Proposition 4.1, we immediately obtain the following corollary (for i below take $F \equiv 1$).

Corollary 4.1 *(Convergence of the marginal distributions of θ and weak convergence of the* *marginal distributions of S_k).*

$$i) \quad \lim_{\substack{H\to\infty\\ \delta\to 0}} \bar{p}(\theta|y^k) = \lim_{\substack{H\to\infty\\ \delta\to 0}} \sum_{h=1}^{4H^2+8} \bar{p}(\underline{r}^h,\theta|y^k)$$
$$= \int dp(S_k,\theta|y^k) = p(\theta|y^k);$$
$$ii) \quad \lim_{\substack{H\to\infty\\ \delta\to 0}} \sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(\underline{r}^h|y^k) = \lim_{\substack{H\to\infty\\ \delta\to 0}} \sum_{\theta}\sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(\underline{r}^h,\theta|y^k)$$
$$= \sum_{\theta}\int F(S_k)dp(S_k,\theta|y^k) = \int F(S_k)dp(S_k|y^k).$$

Remark:

Since power functions are not uniformly continuous nor bounded, Proposition 4.1 and Corollary 4.1 would not allow us to obtain convergence of the conditional moments. We can, however, obtain their convergence by truncating them with an arbitrarily large truncation factor.

5 Empirical Analysis

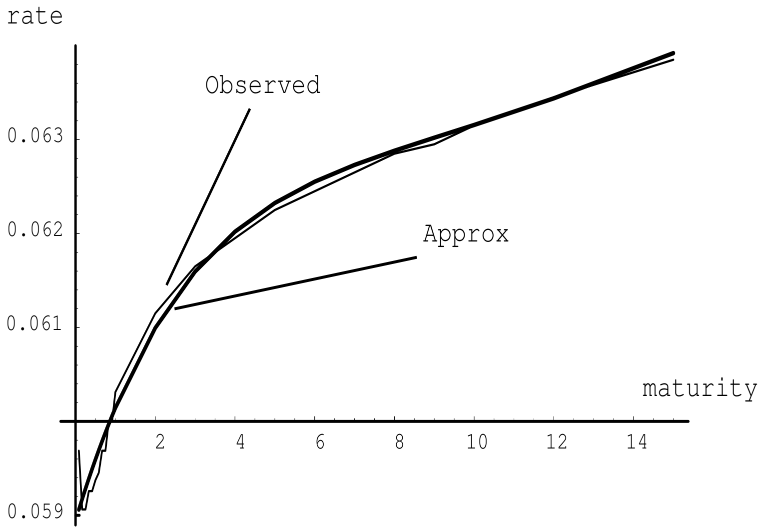


Figure 2: The initial yield curve and its interpolant on 1st December 1997.

We take the daily 1 year US swap rates from Datastream for the period 1st December 1997 to 1st December 1998, as $L_x(t)$ with $x = 1$. The initial yield curve (thin line in figure 2) is formed using the 1 to 11 months US LIBOR rates and the 1 to 15 years US swap rates observed on 1st December 1997. This curve is approximated by the function $r(0,t) = a_0 + (b_0 + b_1 t + b_2 t^2)e^{-b_3 t}$ where a_0, b_0, b_1, b_2 and b_3 are obtained by least squares optimisation.

The parameters $\theta = (\lambda, \sigma_0, \delta)$ are constrained to be within the set $[-1.9, 3] \times [0.01, 0.2] \times [0.05, 2]$ and $r = (r(t, 0), r(t, x))$ to be within $[0.02, 0.09] \times [0.02, 0.09]$. The parameter space is discretised into $50 \times 10 \times 50$ cells to give a total of 28611 combinations for θ, which are stored in a vector such that the first element is $(-1.9, 0.01, 0.05)$, increments are made to δ, σ_0, then λ, and the last element is $(3, 0.2, 2)$. The vector r is discretised as described in section 4.

In figure 3, the distribution $\bar{p}(S_k = r^h|y_k)$ is plotted against $r(t, 0)$ and $r(t, x)$, $x = 1$. The peak of the distribution lies approximately on the circle with radius 0.045. This can be seen more clearly in the contour plot in figure 4.

Figure 5 shows a 3D contour plot of the distribution $\bar{p}(\theta|y_k)$ against $\lambda \in [-1.9, 3], \sigma \in [0.01, 0.2], \delta \in [0.05, 2]$. The colored region is the region where the probability is greater than zero. Figure 6 shows the distribution against the theta number. The maximum occurs at number 10679 which corresponds to $\lambda = -0.038, \sigma = 0.01, \delta = 0.791$. For $\sigma = 0.01$ the distribution is shown in figures 7 and 8. A contour plot is shown in figure 9.

It is interesting to note that the value of δ (0.791) at which the maximum occurs is not far from the value of 0.7079 obtained by Sun (2003) using weekly US T-bill rates and maximum likelihood estimation of a simple discretisation of the standard stochastic differential equation for the instantaneous spot rate.

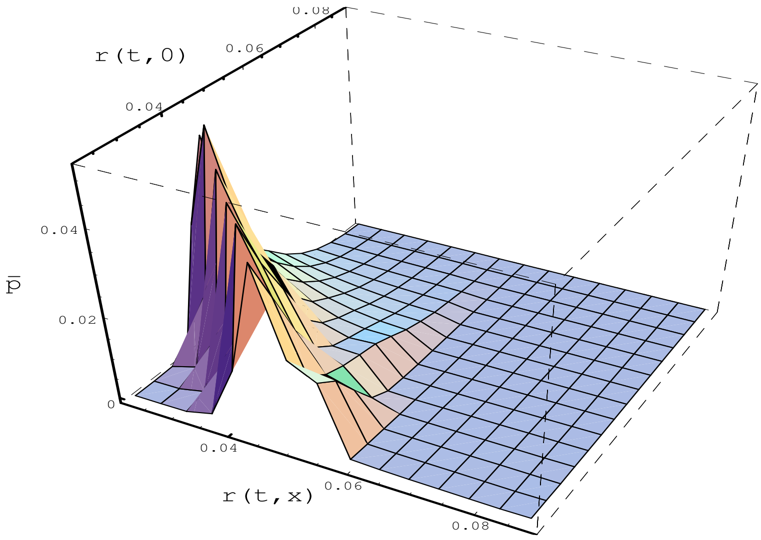


Figure 3: Plot of $\bar{p}(S_k = r^h | y_k)$ against $r(t, 0)$ and $r(t, x)$. The range for these two rates was chosen to be $[0.02, 0.09]$.

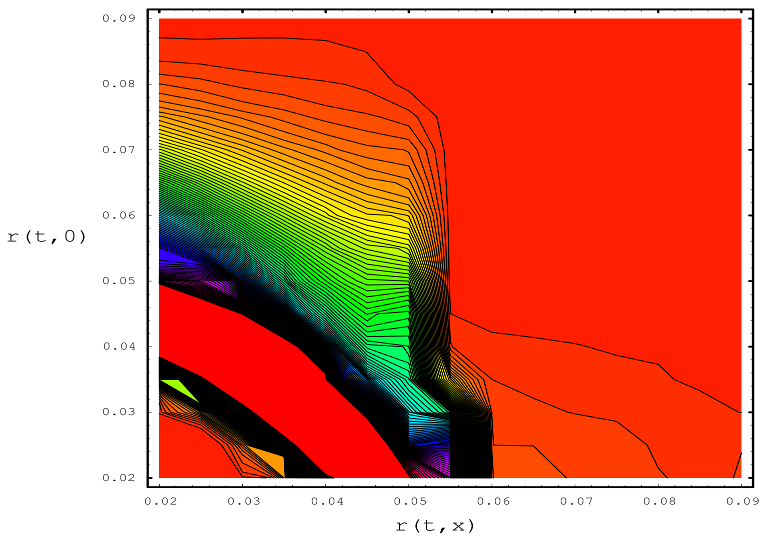


Figure 4: Contour plot of $\bar{p}(S_k = r^h | y_k)$ shown in the previous figure.

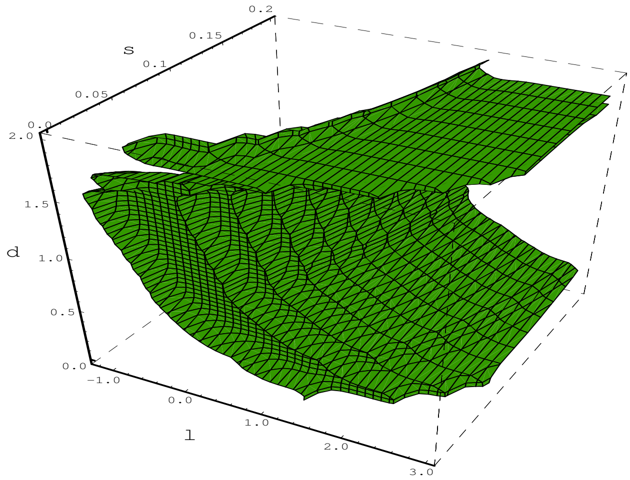


Figure 5: 3D contour plot of $\bar{p}(\theta|y_k)$. In the green region the probability is nonzero.

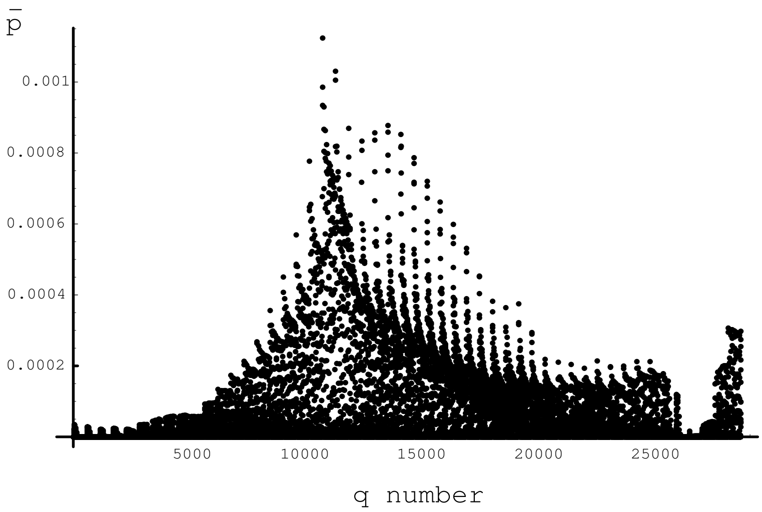


Figure 6: Plot of $\bar{p}(\theta|y_k)$ versus theta number.

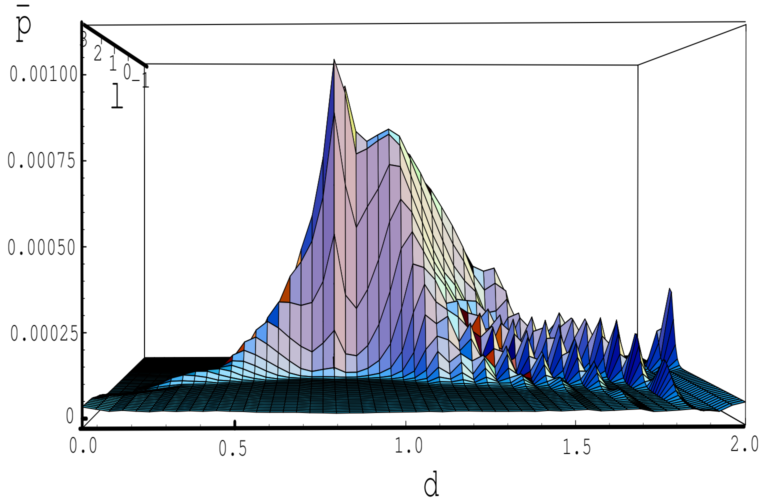


Figure 7: Plot of $\bar{p}(\theta|y_k)$ viewed parallel to the λ axis. The peak lies at $\delta = 0.7$, but the centre is larger than that value.

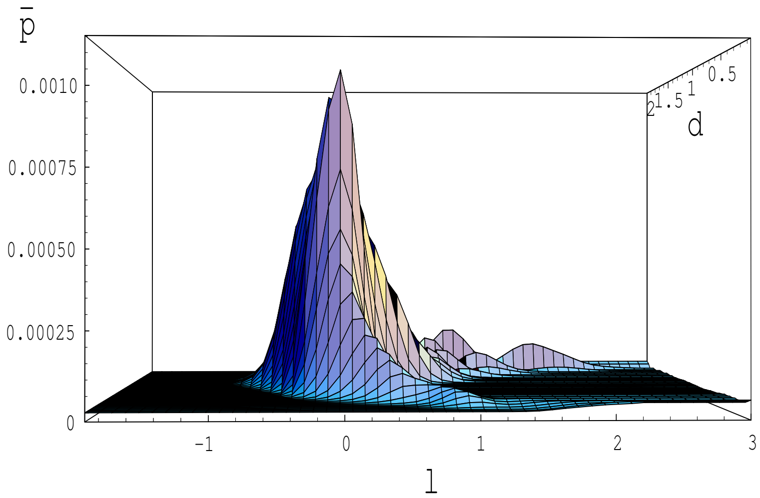


Figure 8: Plot of $\bar{p}(\theta|y_k)$ viewed along the $delta$ axis.

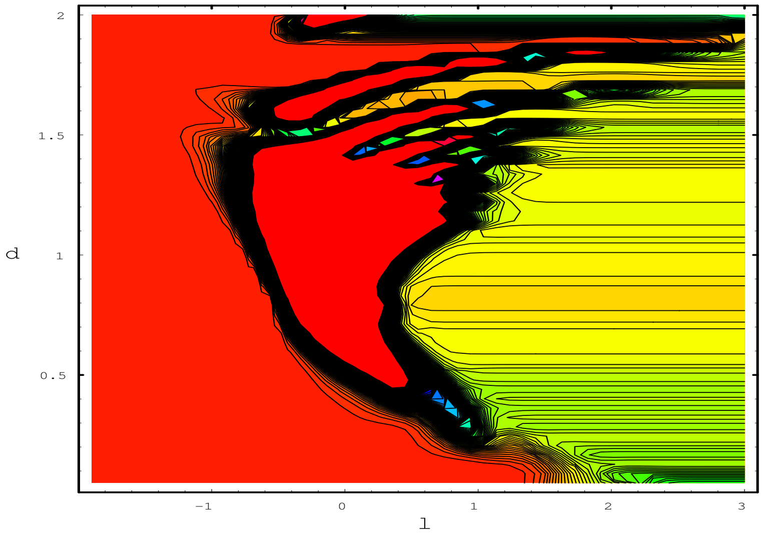


Figure 9: Contour plot of $\bar{p}(\theta|y_k)$.

6 Conclusion

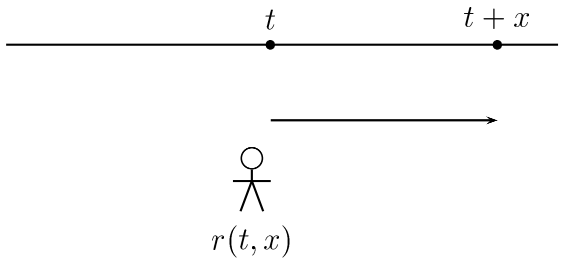
We have derived the risk-neutral dynamics for unobserved factors upon which pure discount bond prices depend within the Heath-Jarrow-Morton framework using a certain forward rate volatility specification. We have then used the link between LIBOR rates, forward rates and pure discount bond prices to obtain the corresponding dynamics for LIBOR rates. The overall stochastic dynamic system can then be treated as a partially observed system with changes in the LIBOR rates being the observations. We have considered a discretised version of the model and developed a dynamic Bayesian updating algorithm to obtain the distribution of the model parameters conditional on the observed market LIBOR rates. The algorithm has been applied to some U.S. data and gives a model fit that seems consistent with some of the traditional econometric studies.

In addition to avoiding the use of proxy variables for the instantaneous spot rate of interest, the methodology proposed here has the advantage that a number of available discretely compounded rates may be used as the observed quantities. In this way one could obtain the volatility for the instantaneous spot rate most consistent with a set of discretely compounded LIBOR rates whose maturities are of most relevance to the application at hand.

Future research needs to relax some of our restrictive assumptions, in particular the non-stochasticity of the market price of interest rate risk. More work also needs to be done on statistical diagnostics to assess the goodness-of-fit of the estimated models.

7 Appendix 1: BM Dynamics - Forward Rate Dependent Volatility Function

Consider the HJM model within the BM parameterisation. We have the forward rate $r(t, x)$.



Under the risk neutral measure $\widetilde{\mathbb{P}}$ it satisfies the stochastic integal equation

$$r(t,x) = r(0,t+x) + \int_0^t \sigma(s,t+x)\bar{\sigma}(s,t+x)ds + \int_0^t \sigma(s,t+x)d\widetilde{w}(s), \tag{52}$$

where

$$\bar{\sigma}(s,t+x) = \int_s^{t+x} \sigma(s,u)du. \tag{53}$$

From(52) with $x = 0$ we obtain the stochastic integal equation for the instantaneous spot rate $r(t)(= r(t,0))$

$$r(t) = r(0,t) + \int_0^t \sigma(v,t)\bar{\sigma}(v,t)dv + \int_0^t \sigma(v,t)d\widetilde{w}(v). \tag{54}$$

Here we consider volatility functions of the form

$$\sigma(t,u) = g\left(r(t,x_1), ..., r(t,x_n)\right) e^{-\lambda(u-t)}. \tag{55}$$

The dynamics for each state variable $r(t,x_i)(i = 1,...,n)$ is obtained by setting $x = x_i$ in (52). We note that with the specification (55), the expression (53) for $\bar{\sigma}(s,t)$ becomes

$$\bar{\sigma}(s,t) = \frac{(1 - e^{-\lambda(t-s)})}{\lambda} g\left(r(s,x_1), ..., r(s,x_n)\right). \tag{56}$$

To ease the notation we set

$$g\left(\vec{r}(s,\cdot)\right) \equiv g\left(r(s,x_1), ..., r(s,x_n)\right), \tag{57}$$

so that we can write

$$\sigma(t,u) = g\left(\vec{r}(t,\cdot)\right) e^{-\lambda(u-t)},$$

and

$$\bar{\sigma}(t,u) = g\left(\vec{r}(t,\cdot)\right)\frac{1-e^{\lambda(u-t)}}{\lambda}.$$

With these various notations the stochastic integal equation for $r(t,x_i)$ becomes

$$\begin{aligned} r(t,x_i) = r(0,t+x_i) + \int_0^t g^2\left(\vec{r},(s,\cdot)\right)e^{-\lambda(t+x_i-s)}\frac{1-e^{-\lambda(t+x_i-s)}}{\lambda}ds \\ + \int_0^t g\left(\vec{r}(s,\cdot)\right)e^{-\lambda(t+x_i-s)}d\widetilde{w}(s), \end{aligned} \tag{58}$$

for $i = 1.2, ..., n.$

Note that in terms of the Chiarella and Kwon (2002) notation we may write

$$\sigma(t,u) = c_{11}(t)\sigma_{11}(u), \tag{59}$$

where

$$\sigma_{11}(u) = e^{-\lambda u}, \quad c_{11}(t) = e^{\lambda t}g\left(\vec{r}(t,\cdot)\right), \tag{60}$$

and so

$$\bar{\sigma}_{11}(x) = \frac{1-e^{-\lambda x}}{\lambda}. \tag{61}$$

The naturally arising subsidiary variables needed to Markovianise the dynamics then turn out to be

$$\psi_1'(t) = \int_o^t e^{\lambda s}g\left(\vec{r}(s,\cdot)\right)d\widetilde{w}(s) - \int_o^t g\left(\vec{r}(s,\cdot)\right)^2 e^{2\lambda s}\frac{\left(1-e^{-\lambda s}\right)}{\lambda}ds, \tag{62}$$

$$\varphi_{11}'(t) = \int_o^t c_{11}(s)^2 ds = \int_o^t g\left(\vec{r}(s,\cdot)\right)^2 e^{2\lambda s}ds. \tag{63}$$

The quantities $\psi_1'(t)$, $\varphi_{11}'(t)$ are both stochastic, so in terms of Chiarella and Kwon (2002) notation the stochastic differential equation for $r(t,x)$ is written

$$r(t,x) = r(0,t+x) + \sigma_{11}(t+x)\psi_1'(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\varphi_{11}'. \tag{64}$$

By choosing any two values for x e.g. x_1, x_2, we have two equations for $\psi_1^{'}(t)$ and $\varphi_{11}^{'}(t)$ in terms of $r(t,x_1)$ and $r(t,x_2)$.

Thus we have the system (henceforth we write $\psi(t), \varphi(t)$)

$$\begin{bmatrix} \sigma_{11}(t+x_1) & \sigma_{11}(t+x_1)\bar{\sigma}_{11}(t+x_1) \\ \sigma_{11}(t+x_2) & \sigma_{11}(t+x_2)\bar{\sigma}_{11}(t+x_2) \end{bmatrix}\begin{bmatrix} \psi(t) \\ \varphi(t) \end{bmatrix} = \begin{bmatrix} r(t,x_1) - r(0,t+x_1) \\ r(t,x_2) - r(0,t+x_2) \end{bmatrix},$$

so that

$$\begin{bmatrix} \psi(t) \\ \varphi(t) \end{bmatrix} = \begin{bmatrix} \Delta_{11}(t) & \Delta_{12}(t) \\ \Delta_{21}(t) & \Delta_{22}(t) \end{bmatrix} \begin{bmatrix} r(t,x_1) - r(0,t+x_1) \\ r(t,x_2) - r(0,t+x_2) \end{bmatrix}, \tag{65}$$

where

$$\boxed{\begin{aligned} &\Delta_{11}(t) = \sigma_{11}(t+x_2)\bar{\sigma}_{11}(t+x_2)/\Delta, \quad \Delta_{12}(t) = -\sigma_{11}(t+x_1)\bar{\sigma}_{11}(t+x_1)/\Delta, \\ &\Delta_{21}(t) = -\sigma_{11}(t+x_2)/\Delta, \quad \Delta_{22}(t) = \sigma_{11}(t+x_1)/\Delta, \\ &\Delta = \sigma_{11}(t+x_1)\sigma_{11}(t+x_2)\left(\bar{\sigma}_{11}(t+x_2) - \bar{\sigma}_{11}(t+x_1)\right). \end{aligned} \quad (66)}$$

Rearranging (65) we can express $\psi(t)$ and $\varphi(t)$ as

$$\psi(t) = \Delta_{11}(t)r(t,x_1) + \Delta_{12}(t)r(t,x_2) - \delta_1(t), \tag{67}$$
$$\varphi(t) = \Delta_{21}(t)r(t,x_1) + \Delta_{22}(t)r(t,x_2) - \delta_2(t), \tag{68}$$

where

$$\begin{aligned} \delta_1(t) &= \Delta_{11}(t)r(0,t+x_1) + \Delta_{12}(t)r(0,t+x_2), \\ \delta_2(t) &= \Delta_{21}(t)r(0,t+x_1) + \Delta_{22}(t)r(0,t+x_2). \end{aligned}$$

Thus the forward rate of any tenor can be expressed in terms of $r(t,x_1)$ and $r(t,x_2)$ by substituting (67) and (68) into (64) i.e.

$$\begin{aligned} r(t,x) = r(0,t+x) &- \sigma_{11}(t+x)\delta_1(t) - \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\delta_2(t) \\ &+ \left[\sigma_{11}(t+x)\Delta_{11}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t)\right] r(t,x_1) \\ &+ \left[\sigma_{11}(t+x)\Delta_{12}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{22}(t)\right] r(t,x_2). \end{aligned} \tag{69}$$

In terms of equation (22) of Chiarella and Kwon (2003) we have

$$\boxed{\begin{aligned} b_0(t,x) &= r(0,t+x) - \sigma_{11}(t+x)\delta_1(t) - \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\delta_2(t), \\ b_1(t,x) &= \sigma_{11}(t+x)\Delta_{11}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t), \\ b_2(t,x) &= \sigma_{11}(t+x)\Delta_{12}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{22}(t). \end{aligned} \quad (70)}$$

Thus (69) can be written

$$r(t,x) = b_0(t,x) + \sum_{i=1}^{2} b_i(t,x)r(t,x_i). \tag{71}$$

We can then use equation (25) from Chiarella and Kwon (2003) to obtain the expression for the BM bond price

$$\boxed{b(t,x) = exp(-[\bar{b}_o(t,x) + \sum_{i=1}^{2} \bar{b}_i(t,x)r(t,x_i)]), \quad (72)}$$

where

$$\bar{b}_i(t,x) = \int_0^x b_i(t,u)du, \qquad (i = 0,1,2). \tag{73}$$

The $x-$ tenor LIBOR rate is related to $b(t,x)$ via

$$L_x(t) = \frac{1}{x}\left(\frac{1}{b(t,x)} - 1\right). \tag{74}$$

We treat (74) as the observation equation with the state variables $r(t,x_1),...,r(t,x_n)$ being driven the stochastic differential equation system (see Chiarella and Kwon (2002) just below equation (23)

$$\begin{aligned} dr(t,x_k) = &\left[b_0^{'}(t,x_k) + \sum_{i=1}^{2} b_i^{'}(t,x_k)r(t,x_i) + \sigma(t,t+x_k)\bar{\sigma}(t,t+x_k)\right] dt \\ &+ e^{-\lambda x_k} g(r(t,x_1),...,r(t,x_n))d\widetilde{w}(t), \qquad (k = 1,2,...,n). \end{aligned} \tag{75}$$

Note that in (75)

$$b_i^{'}(t,x) \equiv \frac{\partial b_i(t,x)}{\partial x}. \tag{76}$$

From the particular specification in (59) we have

$$\sigma_{11}(t+x) = e^{-\lambda(t+x)}, \tag{77}$$

so that

$$\frac{\partial \sigma_{11}(t+x)}{\partial x} = -\lambda e^{-\lambda(t+x)} = -\lambda\sigma_{11}(t+x), \tag{78}$$

and

$$\bar{\sigma}_{11}(t+x) = \frac{1-e^{-\lambda(t+x)}}{\lambda} = \frac{1-\sigma_{11}(t+x)}{\lambda}, \tag{79}$$

$$\bar{\sigma}_{11}'(t+x) = -\frac{\sigma_{11}'(t+x)}{\lambda} = \sigma_{11}(t+x). \tag{80}$$

From equation (70) we calculate

$$\begin{aligned} b_0'(t,x) &= r'(0,t+x) - \sigma_{11}'(t+x)\left[\delta_1(t) + \bar{\sigma}_{11}(t+x)\delta_2(t)\right] - \bar{\sigma}_{11}'(t+x)\sigma_{11}(t+x)\delta_2(t) \\ &= r'(0,t+x) + \lambda\sigma_{11}(t+x)\left[\delta_1(t) + \bar{\sigma}_{11}(t+x)\delta_2(t)\right] - \sigma_{11}^2(t+x)\delta_2(t) \\ &= r'(0,t+x) + \sigma_{11}(t+x)\left[\lambda\delta_1(t) + (1-2\sigma_{11}(t+x))\delta_2(t)\right], \end{aligned} \tag{81}$$

$$\begin{aligned} b_1'(t,x) &= \sigma_{11}'(t+x)\Delta_{11}(t) + \sigma_{11}'(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t) + \bar{\sigma}_{11}'(t+x)\sigma_{11}(t+x)\Delta_{21}(t) \\ &= -\lambda\sigma_{11}(t+x)[\Delta_{11}(t) + \bar{\sigma}_{11}(t+x)\Delta_{21}(t)] + \sigma_{11}^2(t+x)\Delta_{21}(t) \\ &= \sigma_{11}(t+x)[-\lambda\Delta_{11}(t) - \lambda\bar{\sigma}_{11}(t+x)\Delta_{21}(t) + \sigma_{11}(t+x)\Delta_{21}(t)] \\ &= \sigma_{11}(t+x)[-\lambda\Delta_{11}(t) + (\sigma_{11}(t+x)-1)\Delta_{21}(t) + \sigma_{11}(t+x)\Delta_{21}(t)] \\ &= \sigma_{11}(t+x)[(2\sigma_{11}(t+x)-1)\Delta_{21}(t) - \lambda\Delta_{11}(t)], \end{aligned} \tag{82}$$

and

$$b_2'(t,x) = \sigma_{11}(t+x)[(2\sigma_{11}(t+x)-1)\Delta_{22}(t) - \lambda\Delta_{12}(t)]. \tag{83}$$

To operationalise equation (72) we calculate according to equation (73) the quantities [2]

$$\begin{aligned}
\bar{b}_0(t,x) &= \int_0^x \left[r(0,t+u) - \sigma_{11}(t+u)\delta_1(t) - \sigma_{11}(t+u)\bar{\sigma}_{11}(t+u)\delta_2(t)\right] du \\
&= \bar{r}(0,t+x) - \delta_1(t)\int_0^x e^{-\lambda(t+u)}du - \delta_2(t)\int_0^x \frac{e^{-\lambda(t+u)} - e^{-2\lambda(t+u)}}{\lambda}du \\
&= \bar{r}(0,t+x) + \frac{\lambda\delta_1(t)+\delta_2(t)}{\lambda^2}(\sigma_{11}(t+x)-\sigma_{11}(t)) + \frac{\delta_2(t)}{2\lambda^2}(\sigma_{11}^2(t) - \sigma_{11}^2(t+x)),
\end{aligned} \tag{84}$$

$$\begin{aligned}
\bar{b}_1(t,x) &= \int_0^x \sigma_{11}(t+u)\Delta_{11}(t) + \sigma_{11}(t+u)\bar{\sigma}_{11}(t+u)\Delta_{21}(t)du \\
&= \int_0^x \sigma_{11}(t+u)\Delta_{11}(t) + \frac{\sigma_{11}(t+u)-\sigma_{11}^2(t+u)}{\lambda}\Delta_{21}(t)du \\
&= \frac{\lambda\Delta_{11}(t)+\Delta_{21}(t)}{\lambda^2}(\sigma_{11}(t)-\sigma_{11}(t+x)) + \frac{\Delta_{21}(t)}{2\lambda^2}(\sigma_{11}^2(t+x)-\sigma_{11}^2(t)),
\end{aligned} \tag{85}$$

and

$$\bar{b}_2(t,x) = \frac{\lambda\Delta_{12}(t)+\Delta_{22}(t)}{\lambda^2}(\sigma_{11}(t)-\sigma_{11}(t+x)) + \frac{\Delta_{22}(t)}{2\lambda^2}(\sigma_{11}^2(t+x)-\sigma_{11}^2(t)). \tag{86}$$

For our application $x_1 = 0$ and $x_2 = x$ so in equation (66) we set $x_1 = 0, x_2 = x$, to obtain

$$\begin{aligned}
\Delta_{11}(t) &= e^{2\lambda t}\frac{1-e^{-\lambda(t+x)}}{1-e^{\lambda x}}, \quad \Delta_{12}(t) = -e^{\lambda(2t+x)}\frac{1-e^{-\lambda t}}{1-e^{-\lambda x}}, \\
\Delta_{21}(t) &= -e^{2\lambda t}\frac{\lambda}{1-e^{-\lambda x}}, \quad \Delta_{22}(t) = e^{\lambda(2t+x)}\frac{\lambda}{1-e^{-\lambda x}}.
\end{aligned}$$

From equations (12) and (53) we find that

$$\bar{\sigma}(t,t+x) = \frac{\sigma_0 r^\delta}{\lambda}(1-e^{-\lambda x}).$$

[2]Note that $\bar{r}(0,t+x) \equiv \int_0^x r(0,t+u)du$

8 Appendix 2: Proof of Proposition 4.1.

(The proof is an adaptation to the present situation of the one of Theorem 4.1 in Bhar, Chiarella and Runggaldier (2000).)

In the proof we shall use S_k^H to denote, for the generic period $k = k\Delta t$, the random variable previously denoted by $\bar{S}_k$ that takes the values $\underline{r}^h$ $(h = 1, \cdots, 4H^2 + 8)$ of the representative elements in the discretization defined in this section and has distribution $\bar{p}(\bar{S}_k, \theta | y^k)$ (see (44)) for a given value of θ. The proof then amounts to showing that, for any uniformly continuous function F,

$$\lim_{\substack{H \to \infty \\ \delta \to 0}} E_\theta\{F(S_k^H)|y^k\} = E_\theta\{F(S_k)|y^k\}. \tag{87}$$

We proceed by induction on k. For $k = 0$ we have by construction that

$$\lim_{\substack{H \to \infty \\ \delta \to 0}} E_\theta\{F(S_0^H)\} = E_\theta\{F(S_0)\}. \tag{88}$$

Assume then that the statement is true for k. In period $(k+1)\Delta t$ consider now (see (44))

$$\begin{aligned}
E_\theta\{F(S_{k+1}^H)|y^{k+1}\} &= \sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(S_{k+1}^H = \underline{r}^h|y^{k+1}) \\
\propto \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta) &\sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)}\bar{p}(\underline{r}^i, \theta|y^k) \\
= \sum_{i=1}^{4H^2+8} \phi_F^H(\underline{r}^i, \theta, y_{k+1})\bar{p}(\underline{r}^i, \theta|y^k) &= E\{\phi_F^H(S_k^H, \theta, y_{k+1})|y^k\}.
\end{aligned} \tag{89}$$

With

$$\begin{aligned}
\phi_F^H(\underline{r}^i, \theta, y_{k+1}) &= \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta)P_{[i,h]}^{(k,\theta)} \\
= \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta) &\int_{R^h} dp_\theta(S_{k+1}|S_k^H = \underline{r}^i)
\end{aligned} \tag{90}$$

where the rightmost equality follows from (43). Notice that

$$E_\theta\{F(S_{k+1}^H)|y^{k+1}\} = \frac{E_\theta\{\phi_F^H(S_k^H, \theta, y_{k+1}|y^k\}}{E_\theta\{\phi_1^H(S_k^H, \theta, y_{k+1})|y^k\}}. \tag{91}$$

On the other hand, according to (40) we have

$$\begin{aligned}
E_\theta\{F(S_{k+1})|y^{k+1}\} &= \int F(S_{k+1})dp(S_{k+1}, \theta|y^{k+1}) \\
\propto \int F(S_{k+1})p(y_{k+1}|S_{k+1}, \theta) &\int dp_\theta(S_{k+1}|S_k)dp(S_k, \theta|y^k) \\
= \int \phi_F(S_k, \theta, y_{k+1})dp(S_k, \theta|y^k).&
\end{aligned} \tag{92}$$

With

$$\phi_F(S_k,\theta,y_{k+1}) = \int F(S_{k+1})p(y_{k+1}|S_{k+1},\theta)dp_\theta(S_{k+1}|S_k) \tag{93}$$

and notice that, here too,

$$E_\theta\{F(S_{k+1})|y^{k+1}\} = \frac{E_\theta\{\phi_F(S_k,\theta,y_{k+1}|y^k\}}{E_\theta\{\phi_1(S_k,\theta,y_{k+1})|y^k\}}. \tag{94}$$

Given our assumptions, the function $F(S)p(y|S,\theta)$ with $p(y|S,\theta)$ as in (38) is uniformly continuous in S for all values of y and θ. Notice, in fact, that $p(y|S,\theta)$ is continuous in S and its limit for S going to infinity is, uniformly in (y,θ) equal to zero. Given $\eta > 0$, there is thus H_F^o (depending on F and the parameters in the model such as x and Λ_k) such that, for $H > H_F^o$ and $\delta < (H_F^o)^{-1}$, one has in any period $k\Delta t$ and for all values of S_k,θ,y_{k+1},

$$\begin{aligned}
&|\phi_F^H(S_k,\theta,y_{k+1}) - \phi_F(S_k,\theta,y_{k+1})| \\
\leq &\sum_{h=1}^{4H^2+8} \int_{R^h} |F(\underline{r}^h)p(y_{k+1}|\underline{r}^h,\theta) - F(S_{k+1})p(y_{k+1}|S_{k+1},\theta)dp_\theta(S_{k+1}|S_k) \\
\leq &\quad \eta \sum_{h=1}^{4H^2+8} \int dp_\theta(S_{k+1}|S_k) = \eta.
\end{aligned} \tag{95}$$

To complete the induction step we have to show that $\phi_F(S;\theta,y)$ is, for all values of (θ,y), uniformly continuous in S. In fact, if this is the case, then, given $\eta > 0$, there exists $H_F^1(y^{k+1},\theta)$ (depending on F and the parameters in the model) such that, for $H > H_F^1$ and $\delta < (H_F^1)^{-1}$,

$$|E_\theta\{\phi_F(S_k^H,\theta,y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k,\theta,y_{k+1})|y^k\} \leq \eta. \tag{96}$$

From (95) and (96) it then follows that

$$\begin{aligned}
&|E_\theta\{\phi_F^H(S_k^H;\theta,y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k;\theta,y_{k+1})|y^k\}| \\
\leq &|E_\theta\{\phi_F^H(S_k^H;\theta,y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k^H;\theta,y_{k+1})|y^k\}| \\
&+ |E_\theta\{\phi_F(S_k^H;\theta,y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k;\theta,y_{k+1})| \leq 2\eta
\end{aligned} \tag{97}$$

for $H > H_F := \max\{H_F^o, H_F^1(y^{k+1},\theta)\}$ and $\delta < (H_F)^{-1}$. Combining (97) with (91) and (94) allows to complete the induction step and thus the proof of the proposition.

It remains thus to show that $\phi_F(S;\theta,y)$ is uniformly continuous in S for each pair (θ,y).

To this effect notice that, by (36) as well as (38) and (39) (see also (93)) we have that

$$\begin{aligned}
\phi_F(S_k,\theta,y^{k+1}) &= \int F(S_{k+1}^1,S_{k+1}^2)p(y_{k+1}|S_{k+1}^1,S_{k+1}^2,\theta)\cdot dp_\theta(S_{k+1}^1,S_{k+1}^2|S_k^1,S_k^2) \\
&= \int \bar{F}(S_{k+1}^1,S_k^1,S_k^2)\bar{p}(y_{k+1}|S_{k+1}^1,S_k^1,S_k^2,\theta)\cdot p_\theta(S_{k+1}^1|S_k^1,S_k^2)dS_{k+1}^1,
\end{aligned} \tag{98}$$

where

$$\bar{F}(S_{k+1}^1,S_k^1,S_k^2) := F(S_{k+1}^1,\alpha S_{k+1}^1 + \beta(S_k^1,S_k^2)) \tag{99}$$

and $p_\theta(S^1_{k+1}|S^1_k,S^2_k)$ coincides with the Gaussian density $p_\theta(S^1_{k+1}|S_k)$ in (37).

Since (see comment after (35)) $\beta(S^1_k,S^2_k)$ is uniformly continuous we have that $\bar{F}(S^1_{k+1},S^1_k,S^2_k)$ is bounded and uniformly continuous in $S_k=(S^1_k,S^2_k)$.

We next have from (37) with (35) and (22) that

$$\begin{aligned}
&p_\theta(S^1_{k+1}|S_k)=\frac{1}{\sqrt{2\pi}g(r(k,0))}\cdot\exp\left\{-\frac{1}{2g^2(r(k,0))}(S^1_{k+1}-S^1_k-R(r(k,x),r(k,0)))^2\right\}\\
&:=\phi_1(S_k)\cdot\phi_2(S_k;S^1_{k+1})
\end{aligned}\tag{100}$$

where (see (11) and (24)) $\phi_1(S_k)$ and $\phi_2(S_k;S^1_{k+1})$ are bounded functions that are uniformly continuous in S_k. Concerning $\phi_2(S_k;S^1_{k+1})$ notice in fact that it is continuous and the limit for S_k going to infinity is, uniformly in S^1_{k+1}, equal to zero. Being a product of two bounded and uniformly continuous functions, $p_\theta(S^1_{k+1}|S_k)$ is thus bounded and uniformly continuous in S_k.

Coming to $\bar{p}(y_{k+1}|S^1_{k+1},S_k;\theta)$ defined in (39), it is immediately seen that, as a function of S_k, it is bounded uniformly in (y_{k+1},S_{k+1},θ). Since it is continuous in S_k and the limit for S_k going to infinity is zero uniformly in (y_{k+1},S_{k+1},θ), it is also a uniformly continuous function of S_k. Notice, furthermore, that the integral with respect to S^1_{k+1} is equal to the constant $\gamma=\bar{b}_1((k+1)\Delta t,x)+\bar{b}_2((k+1)\Delta t,x)\alpha$.

Take now S_k, $\tilde{S}_k\in\mathbb{R}^2$. By (98), the previously shown uniform continuity properties and uniform integrability in S^1_{k+1} of $p_\theta(S^1_{k+1}|S^1_k,S^2_k)$ and $\bar{p}(y_{k+1}|S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k,\theta)$ we have that, given $\eta>0$, one can choose $\delta>0$ such that for $\|S_k-\tilde{S}_k\|<\delta$ it results that

$$\begin{aligned}
&|\phi_F(S_k,\theta,y_{k+1})-\phi_F(\tilde{S}_k,\theta,y_{k+1}|\\
\leq&\int|\bar{F}(S^1_{k+1},S^1_k,S^2_k)\bar{p}(y_{k+1}|S^1_{k+1},S^1_k,S^2_k,\theta)\\
&\quad-\bar{F}(S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k)\bar{p}(y_{k+1}|S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k,\theta)|p_\theta(S^1_{k+1}|S^1_k,S^2_k)dS^1_{k+1}\\
&\quad+\int\bar{F}(S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k)|p_\theta(S^1_{k+1}|S^1_k,S^2_k)-p_\theta(S^1_{k+1}|\tilde{S}^1_k,\tilde{S}^2_k)|\bar{p}(y_{k+1}|S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k,\theta)dS_{k+1}\\
\leq&\eta\int p_\theta(S^1_{k+1}|S^1_k,S^2_k)dS^1_{k+1}+\frac{\eta}{\gamma}\int\bar{p}(y_{k+1}|S^1_{k+1},\tilde{S}^1_k,\tilde{S}^2_k,\theta)dS^1_{k+1}=2\eta.
\end{aligned}\tag{101}$$

Thus showing the uniform continuity in S of $\phi_F(S,\theta,y)$ for each pair (θ,y) and with it also the induction step.

■

9 References

ANDERSEN, T. and LUND, J. (1997): Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics* 77, 343-377.

BALI, T.G. (1999): An empirical comparison of continuous time models of the short-term interest rate. *The Journal of Futures Markets* 19, 777-797.

BERGSTROM, A. (1990): *Continuous time econometric modelling*, Oxford University Press.

BHAR, R. and CHIARELLA, C.(1997): Transformation of Heath-Jarrow-Morton models to Markovian systems. *European Jounal of Finance* 3, 1-26.

BHAR, R., CHIARELLA, C. and RUNGGALDIER, W.(2002): Estimation in Models of the Instantaneous Short Term Interest Rate by Use of a Dynamic Bayesian Algorithm in *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann, Eds. K. Sandmann and P. Schönbucher, Springer-Verlag*, 177-195.

BRENNER, R., HARJES, R. and KRONER, K. (1996): Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis* 31, 85-107.

CHAN, K., KAROLYI, A., LONGSTAFF, F. and SANDERS, A. (1992): An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47, 1209-1227.

CHAPMAN, D., LONG, J. and PEARSON, N.(1999): Using proxies for the short rate: when are three months like an instant? *Review of Financial Studies* 12(4), 763-806.

CHIARELLA,C. and KWON,O.K.(2003): Finite Dimensional Affine Realisations of HJM Models in Terms of Forward Rates and Yields. *Review of Derivatives Research* 6(2), 129-155.

CHIARELLA, C., PASQUALI, S. and RUNGGALDIER, W.J. (2001) : On Filtering in Markovian Term Structure Models (an approximation approach), *Advances in Applied Probability*, 33(4), 794-809, 2001.

CHEYETTE, O.(1992): Term structure dynamics and mortgage valuation. *The Journal of Fixed Income* pp. 28-41.

COX, J., INGERSOLL, J. and ROSS, S. (1985): A theory of the term structure of interest rates. *Econometrica* 53, 385-407.

EPISCOPOS A. (2000): Further evidence on alternative continuous time models of the short-term interest rate. *Journal of International Financial Markets, Institutions & Money* 10, 199-212.

HARVEY A.C. (1989): Forecasting, structural time series models and the Kalman filter. Melbourne: Cambridge University Press.

INUI, K. and KIJIMA, M.(1997): A Markovian framework in multi-factor Heath-Jarrow-Morton models. *Journal of Financial and Quantitative Analysis* 33(3), 423-440.

JAZWINSKI, A.H. (1970): Stochastic Process and Filtering Theory. New York: Academic Press.

JIMINEZ, J.C. and OZAKI, T.(2002): An innovation appproach for the estimation of discretely observed stohastic volatility models. 2^{nd} *International Conference on Financial Engineering and Statistical Finance*, Hitotsubashi University, Tokyo, March 18-19, 2002.

KOEDIJK, K. , NISSEN, F., SCHOTMAN, P. and WOLFF, C. (1997): The dynamics of the short-term interest rate volatility reconsidered. *European Finance Review* 1, 105-130.

LO, A.W. (1988): Maximum likelihood estimation of generalized Itô process with discretely sampled data. *Econometric Teory*, 4(2), 231-247.

LONGSTAFF, F. and SCHWARTZ, E. (1992): Interest rate volatility and the term structure of interest rates. *Journal of Finance* 47, 1259-1282.

NOWMAN, K. (1997): Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 52, 1695-1706.

NOWMAN, K. (1998): Continuous-time short term interest rate models. *Applied Financial Economics* 8, 401-407.

RITCHKEN, P. and SANKARASUBRAMANIAN, L.(1995): Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance* 5(1), 55-72.

SANDMANN, K, AND SONDERMANN, D.(1997): A note on the stability of lognormal interest rate models and the pricing of Eurodollar futures. *Mathematical Finance* 7(2),119-125.

SUN, L. (2003): Nonlinear drift and stochastic volatility: An empirical investigation of short-term interest rate models. *The Journal of Financial Research* XXVI(3), 389-404.

VASICEK, O. (1997): An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177-188.